Multicasting with Untrusted Relays: A Noncoherent Secure Network Coding Approach

Ta-Yuan Liu\textsuperscript{1}, Shih-Chun Lin\textsuperscript{2}, and Y.-W. Peter Hong\textsuperscript{1}

\textsuperscript{1}Inst. of Communications Engineering, National Tsing Hua University, Hsinchu, Taiwan
\textsuperscript{2}Dept. of Electronic and Computer Eng., National Taiwan University of Science and Technology, Taipei, Taiwan
Network coding in general improves throughput and reliability.

It is common to assume that all the relays are trustworthy.

However, in practice, some of them may be provided by a third party which cannot be fully trusted.
Multihop Network with Untrusted Relays

- Untrusted (or third party) relays may potentially be compromised by an outside adversary (or an eavesdropper).

- More relays (trusted or not) provides more paths for simultaneous information transfer, but yields higher risk of being eavesdropped.

- Intuitively, one should recruit untrusted relays ONLY when the secrecy capacity can be improved by doing so.

  - Secrecy capacity: Maximum transmission rate without information leakage
Main Contributions

- Exam the impact of untrusted relays in the multihop network system and determine the optimal input signal that maximizes secrecy capacity when untrusted relays are recruited.

- Discuss the untrusted relays recruitment problem based on the secrecy capacity in two different cases:
  - Case 1: All untrusted relays near the destination are compromised with probability $1$.
  - Case 2: Each untrusted relay is compromised with probability $p$. 
The signal transmitted from the source to the first hop of relays is

\[ X \in \mathcal{F}_q^{m \times T} \]

- \( m \) is the number of relays in the first layer,
- \( T \) is packet length, and \( q \) is field size.

**Random linear network coding:** Each relay forwards a linear combination of its received signals with coefficients chosen uniformly over the finite field \( \mathcal{F}_q \).

**Received signal:**
- **Destination:** \( Y = HX \) where \( H \in \mathcal{F}_q^{n \times m} \).
- **Eavesdropper:** \( Z = GX \) where \( G \in \mathcal{F}_q^{e \times m} \).

\( e \): the number of untrusted relays compromised by the eavesdropper.
**Assumptions:**

- **(i)** We assume that, after a sufficient number of hops, the effective channel matrices $H$ and $G$ are i.i.d. uniform in $\mathcal{F}_q$. [Siavoshani & Fragouli ’12]

- **(ii)** $H$ and $G$ are unknown at all nodes (i.e., a noncoherent framework), e.g., when the encoding vector is NOT appended to the network coding packets.

**Special Case:** When NO untrusted relays are recruited, the system model can be reduced as:

- $X_t \in \mathcal{F}_q^{m_t \times T}$
- $Y_t = H_t X_t$
- $H_t \in \mathcal{F}_q^{m_t \times m_t}$

**Diagram:**

[Diagram of network model with nodes S, E, D, and connections showing $X_t$, $Y_t$, and $H_t$]
The secrecy capacity of a degraded channel is

\[ C_s = \max_{p_x} I(X;Y') - I(X;Z'), \quad [\text{Wyner '75}] \]

\[ \max_{V \to X \to Y,Z} I(V;Y) - I(V;Z) \quad [\text{Csiszar & Korner '78}] \]

- \( V \) is a auxiliary variable.
- It is difficult to joint optimize \( V \) and \( X \).

**Equivalent degraded channel**:

- Focus on the case \( n > e \) (if \( n \leq e, C_s = 0 \))

**Original Channel**: \( Y = HX \) \( \quad Z = GX \)

**Equivalent Degraded Channel**: \( Y' = \begin{bmatrix} G \\ H' \end{bmatrix} X \) \( \quad Z' = GX \)

- Equivalent: Secrecy capacity only depend on \( p(Y|X) \) and \( p(Z|X) \).
- Degraded: \( X \to Y' \to Z' \) forms a Markov chain.
Lemma 1([Siavoshani & Fragouli ’12]): The secrecy capacity is given as

\[ C_s = \max_{\Pi_X} I(\Pi_X; \Pi_{Y'}) - I(\Pi_X; \Pi_{Z'}). \]

where \( \Pi_X \) is the subspace which spanned by the row vectors of \( X \).
Moreover, the distribution of optimal input \( \Pi_X^\ast \) is given by

\[ P_{\Pi_X^\ast}(\pi_x) = \alpha_{d_x} \begin{bmatrix} T \\ d_x \end{bmatrix}^{-1} \]

where \( \alpha_{d_x} \triangleq \Pr[\dim(\Pi_X) = d_x] \) is the probability that \( \Pi_X \) is of dimension \( d_x \).

- Only depend on the subspace spanned by the row vectors of input signal \( X \).
- All subspaces of the same dimension occur with equal probability.
Optimization Problem

- **Input optimization problem:**

  \[ C_s = \max_{\alpha} R(\alpha), \quad \text{subject to } ||\alpha||_1 = 1, \]

  where \( R(\alpha) \triangleq I(\Pi^*_X; \Pi_{Y'}) - I(\Pi^*_X; \Pi_{Z'}) \) and the subspace-dimension probabilities \( \alpha \triangleq [\alpha_0, \ldots, \alpha_{\min(m,T)}]^T. \)

- The rate function can be written as

  \[
  R(\alpha) = - \sum_{d_x=0}^{\min(m,T)} \alpha_{d_x} m_{d_x} \log_2 q - \sum_{d_x=0}^{\min(m,T)} \alpha_{d_x} q^{-nd_x} \cdot \sum_{d_y'=0}^{\min(n,d_x)} \psi(n, d_{y'}) \begin{bmatrix} d_x \\ d_{y'} \end{bmatrix} \log_2(f_{Y'}(d_{y'}, \alpha)) \\
  + \sum_{d_x=0}^{\min(m,T)} \alpha_{d_x} e d_x \log_2 q + \sum_{d_x=0}^{\min(e,d_x)} \alpha_{d_x} q^{-ed_x} \cdot \sum_{d_{z'}=0}^{\min(e,d_x)} \psi(e, d_{z'}) \begin{bmatrix} d_x \\ d_{z'} \end{bmatrix} \log_2(f_{Z'}(d_{z'}, \alpha)),
  \]

  - Too complex to derive analytically.
  - Solved using a projection-based gradient descend algorithm.
    - Converge to the optimal solution.
Numerical Result: Secrecy Rate with Different Input Signals

- $T = 20, n = 8, e = 2, q = 7$
Large field size approximation: When field size \( q \gg 1 \), the secrecy capacity can be approximated as

\[
C_s \approx (\min(m_t + m_u, n_t + n_u) - e)(T - \min(m_t + m_u, n_t + n_u)) \log q,
\]

[Siavoshani & Fragouli ’12]

⇒ Special Case (No Untrusted Relays): \( m_u = n_u = e = 0 \).

\[
C \approx \min(m_t, n_t)(T - \min(m_t, n_t)) \log q.
\]

Question: When should we recruit untrusted relays?

- Case I: All untrusted relays near the destination are compromised with probability 1.
- Case II: Each untrusted relay is compromised with probability \( p \).
Case 1: All Untrusted Relays Near the Destination are Compromised

In this case, we assume that the eavesdropper is near the destination so that all $n_u$ untrusted relays in the last hop are compromised.

**Theorem 1:** Let $d_t = m_t - n_t$ and $d_u = m_u - n_u$. When $T > m_t + \max(m_u, n_t)$, untrusted relays should be recruited if $(d_t, d_u)$ satisfies one of the following conditions.

1. $d_t + d_u \leq 0$ and $d_u > \frac{m_t m_u}{T - m_t - m_u}$
2. $d_t + d_u > 0$ and $d_t < \frac{-n_t n_u}{T - n_t - m_t}$. 

\[ e = n_u \]
Recruit Region

- Eavesdropper can obtain \( e = n_u \) dimension.

- **Large** \( d_u \): Recruiting untrusted relays provide more Tx dimension than Rx dimension.

- **Small** \( d_t \): Lack of transmit dimension in the original system.

- When \( T \to \infty \), the recruit region is characterized by \((d_u, d_t)\) only.
Case 2: Each Untrusted Relay is Compromised with Probability $p$

- There is a total of $r_u$ untrusted relays that may be compromised with probability $p$.
  - The number of compromised relays: $e \sim B(r_u, p)$ (Binomial distribution)

- Outage probability: (The probability of no improvement)

$$
P_{out} \triangleq P_r [C_s(e) - C \leq 0] = P_r \left[ e \geq \frac{(k_1 - k_2)(T - k_1 - k_2)}{(T - k_1)} \right].
$$

where $k_1 = \min(m, n)$ and $k_2 = \min(m_u, n_u)$.
Asymptotic Outage Probability

- Suppose that $r_u \to \infty$ and that $m_u = \beta_m r_u$ and $n_u = \beta_n r_u$ for some positive ratio $\beta_m, \beta_n$.

- In this case, $m_t, n_t$ are negligible compared to $r_u$ (and also $m_u$ and $n_u$).

**Theorem 2:** Let us consider a multihop network with parameters $(m_u, n_u, r_u)$. If $m_u = \beta m r_u$ and $n_u = \beta n r_u$ and $T \geq \min(m_u, n_u)$, then

$$P_{out} \to \begin{cases} 
0 & \text{if } p < \beta \\
1 & \text{if } p \geq \beta
\end{cases}$$

as $r_u \to \infty$, where $\beta = \min(\beta_m, \beta_n)$.

- $\beta \cdot r_u$: Dimension provided for the legitimate parts.
- $p \cdot r_u$: Dimension eavesdropped by the eavesdropper.
Conclusions

- Consider a non-coherent multihop network system with the help of untrusted relays which are potentially eavesdropped.

- Determine the optimal input signal when untrusted relays are recruited by a gradient descend algorithm.

- Recruiting untrusted relays problem:
  - Case 1: Determine the recruiting region when all untrusted relays near the destination are compromised.
  - Case 2: Derive the outage probability when each untrusted relay is compromised with probability $p$, and show that when $p$ is less than a threshold, one should recruit.
Thank You for Listening~!